

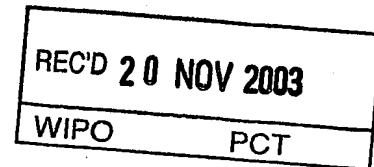
# 证 明

本证明之附件是向本局提交的下列专利申请副本

申 请 日： 2002 11 27

申 请 号： 02 1 50730.9

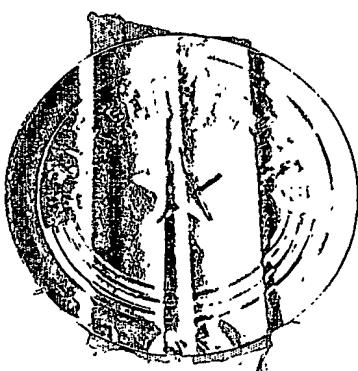
申 请 类 别： 发明



发明创造名称： 人类 1 7 p 1 3 . 3 区域内人肿瘤相关基因 C T 1 2 0 及其编码蛋白

申 请 人： 上海市肿瘤研究所

发明人或设计人： 万大方； 顾健人； 何祥火



**PRIORITY  
DOCUMENT**  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

中华人民共和国  
国家知识产权局局长

王素川

2003 年 10 月 22 日

## 权 利 要 求 书

1. 一种分离的人CT120蛋白多肽，其特征在于，它包括具有SEQ ID NO: 2所示氨基酸序列的多肽，或其保守性变异多肽、或其活性片段、或其活性衍生物。
- 5 2. 如权利要求1所述的多肽，其特征在于，该多肽选自下组：
  - (a) 具有SEQ ID NO: 2氨基酸序列的多肽；
  - (b) 将SEQ ID NO: 2氨基酸序列经过一个或多个氨基酸残基的取代、缺失或添加而形成的，且具有促进NIH/3T3细胞生长功能的由(a)衍生的多肽。
3. 一种分离的多核苷酸，其特征在于，它包含一核苷酸序列，该核苷酸序列与选自下组的一种核苷酸序列有至少85%相同性：
  - (a) 编码如权利要求1和2所述多肽的多核苷酸；
  - (b) 与多核苷酸(a)互补的多核苷酸。
4. 如权利要求3所述的多核苷酸，其特征在于，该多核苷酸编码的多肽具有SEQ ID NO: 2所示的氨基酸序列，或者该多核苷酸具有选自下组的序列：SEQ ID NO: 1中所示的91-861位的编码区序列或1-2145位的全长序列。
- 15 5. 一种载体，其特征在于，它含有权利要求3所述的多核苷酸。
6. 一种遗传工程化的宿主细胞，其特征在于，它是选自下组的一种宿主细胞：
  - (a) 用权利要求5所述的载体转化或转导的宿主细胞；
  - (b) 用权利要求3所述的多核苷酸转化或转导的宿主细胞。
- 20 7. 一种制备多肽的制备方法，其特征在于，该方法包含：
  - (a) 在适合表达蛋白的条件下，培养权利要求6所述的宿主细胞；
  - (b) 从培养物中分离出具有人蛋白CT120活性的多肽。
8. 一种能与权利要求1所述的人CT120蛋白特异性结合的抗体。
9. 一种检测肺细胞是否发生癌变或存在癌变易感性的方法，其特征在于，它25 包括步骤：

检测肺细胞样品中是否有CT120转录本，存在CT120转录本就表示该肺细胞发生癌变或存在癌变易感性；或者

检测肺细胞样品中是否存在CT120蛋白，存在CT120蛋白就表示该肺细胞发生癌变或存在癌变易感性。
- 30 10. 一种检测肺癌的试剂盒，其特征在于，它包括：
  - (1) 特异性扩增人CT120基因的引物对，或
  - (2) 特异性与CT120蛋白结合的抗体。

# 说 明 书

## 人类 17p13.3 区域内人肿瘤相关基因 CT120 及其编码蛋白

### 5 发明领域

本发明属于生物技术领域，具体地说，本发明涉及新的位于人 17 号染色体短臂 1 区 3 带 3 亚带(17p13.3)的编码人肿瘤相关蛋白的多核苷酸，以及此多核苷酸编码的多肽。本发明还涉及此多核苷酸和多肽的用途和制备。

### 10 背景技术

恶性肿瘤的死亡率在我国仅次于心、脑血管疾病名列第二。人们普遍认为肿瘤是多因子，多步骤引起的疾病。

肿瘤的发生与发展实质是一个克隆演化过程。在此过程中伴随一系列细胞核内遗传物质的改变，包括序列改变如点突变，缺失，插入；结构畸变，如大范围缺失，重排，基因扩增。越来越多的证据表明，在克隆演化过程中的不同阶段存在不同基因的激活和/或失活及其复杂的相互作用。因此，分离与鉴定肿瘤相关基因，可以加深人们对肿瘤发生机制的深入理解并有助于对肿瘤的预防、诊断、治疗与预后。

20 肝细胞肝癌(hepatocellular carcinoma, HCC)是一种在亚洲人群中高发的恶性肿瘤，对于 HCC 发生的分子机制，肿瘤生物学家 20 世纪 90 年代初期前后，许多实验室即已陆续注意到 HCC 患者在染色体 17p13.3 区段存在着杂合性丢失(Fujimori et al.Cancer Res.1991,51:89-93;Boige et al,Cancer Res. 1997, 57:1986-1990; Nagai et al,Oncogene,1997,14:2927-2933)；几乎在同一时期，上海市肿瘤研究所的实验室也发现，在中国人群肝癌患者中，在染色体 17p13.3 区段内存在着高频率的染色体杂合性丢失，由此提示在染色体 17p13.3 高频率杂合性缺失区内可能还存在着一个或几个其它的抑癌基因，有别于位于 17p13.1 区的 p53 抑癌基因，在肝癌的发生发展过程中起重要作用。随后，在肝癌患者中，该实验室首先在国际上确定了该杂合性缺失的最小范围为 0.5Mb(Wang et al,Genes Chromosomes&Cancers,2001,31:221-227)。

30 由于癌症是危害人类健康的主要疾病之一。为了有效地治疗和预防肿瘤(如肝癌)，目前人们已越来越关注肿瘤的早期诊断和基因治疗。因此，本领域迫切需要开发研究新的癌症相关的人蛋白及其激动剂/抑制剂。

## 发明内容

本发明的目的是提供一种新的肿瘤相关蛋白-人 CT120 蛋白多肽以及其片段、类似物和衍生物。

5 本发明的另一目的是提供编码这些多肽的多核苷酸。

本发明的另一目的是提供生产这些多肽的方法以及该多肽和编码序列的用途。

在本发明的第一方面，提供了一种分离的人CT120蛋白多肽，它包括具有SEQ  
10 ID NO: 2所示氨基酸序列的多肽，或其保守性变异多肽、或其活性片段、或其活性衍生物。较佳地，该多肽选自下组：(a)具有SEQ ID NO: 2氨基酸序列的多肽；(b)将SEQ ID NO:2氨基酸序列经过一个或多个氨基酸残基的取代、缺失或添加而形成的，且具有促进NIH/3T3细胞生长功能的由(a)衍生的多肽。

在本发明的第二方面，提供了一种分离的多核苷酸，它包含一核苷酸序列，  
15 该核苷酸序列与选自下组的一种核苷酸序列有至少85%相同性：(a)编码如权利要求  
1和2所述多肽的多核苷酸；(b)与多核苷酸(a)互补的多核苷酸。较佳地，该多核苷  
酸编码的多肽具有SEQ ID NO: 2所示的氨基酸序列；更佳地，该多核苷酸具有选自  
下组的序列：SEQ ID NO: 1中所示的编码区序列(第91-861位)或全长序列。

在本发明的第三方面，提供了含有上述多核苷酸的载体，以及被该载体转化  
20 或转导的宿主细胞或者被上述多核苷酸直接转化或转导的宿主细胞。

在本发明的第四方面，提供了制备肿瘤相关 CT120 蛋白活性的多肽的制备方法，  
该方法包含：(a)在适合表达蛋白的条件下，培养上述被转化或转导的宿主细  
胞；(b)从培养物中分离出具有肿瘤相关 CT120 蛋白活性的多肽。

在本发明的第五方面，提供了与上述的肿瘤相关 CT120 蛋白多肽特异性结合  
25 的抗体。还提供了可用于检测的核酸分子，它含有上述的多核苷酸中连续的 20-150  
个核苷酸。

在本发明的第六方面，提供了一种药物组合物，它含有安全有效量的本发明  
的肿瘤相关 CT120 蛋白的拮抗剂(如反义序列或抗体)以及药学上可接受的载体。这些  
药物组合物可治疗癌症以及细胞异常增殖等病症。

30 在本发明的第七方面，提供了一种检测肺细胞是否发生癌变或存在癌变易感性的方法，包括步骤：检测肺细胞样品中是否有CT120转录本，存在CT120转录本就表示该肺细胞发生癌变或存在癌变易感性；或者检测肺细胞样品中是否存在

CT120蛋白，存在CT120打靶就表示该肺细胞发生癌变或存在癌变易感性。

在本发明的第八方面，提供了一种检测肺癌的试剂盒，它包括：(1)特异性扩增人CT120基因的引物对，或(2)特异性与CT120蛋白结合的抗体。

本发明的其它方面由于本文的技术的公开，对本领域的技术人员而言是显而易见的。

### 附图说明

图1显示了CT120与四种同源物的多序列比对结果。

图2显示了CT120的多组织膜片Northern杂交结果。其中，各泳道如下：

10 1. 心；2. 脑；3. 胎盘；4. 肺；5. 肝；6. 骨骼肌；7. 肾；8. 腺。

图3显示了CT120在不同肿瘤组织中的表达(RT-PCR)情况。各泳道如下：

1. SPC-A-1；2. C-33A；3. SMMC-7721；4. BEL-7402；5. SK-OV-3；6. 5637；7. A431；  
8. MCF-7。

图4显示了CT120转染NIH/3T3细胞结果。

15 图5显示了Western印迹检测CT120在稳定转染细胞系中的表达：泳道1-6分别代表6个克隆。

图6显示了免疫组织化学检测CT120在肺癌及癌旁组织中的表达。A 肺癌组织；B 肺癌癌旁组织。

### 20 具体实施方式

在肝癌的研究中，本发明人首先确定了肝癌组织在17p13.3范围内有高频率LOH(60-100%)。最近，通过对肝癌全基因组扫描也证明17p13.3是LOH的最高区域。本发明人对人17号染色体短臂13.3位点的癌相关表达序列(EST)进行了分离和全长克隆。用对应于17p13.3区段内926位点的噬菌体人工染色体(PAC)579号(P579)克隆，通过九倍鸟枪法(shotgun)测序得到其序列，应用计算机分析在其中找到1个代表新基因的EST，通过RACE方法获得全长核苷酸序列和编码的氨基酸，命名为CT120。Northern、Southern杂交等实验证实，CT120在肺癌及癌旁组织中的表达：肺癌细胞中高表达，癌旁组织几乎不表达。这表明CT120与肿瘤相关，体外实验证明对小鼠NIH/3T3细胞具有促进细胞转化功能。因此，CT120基因是一种候选癌基因，可应用于肿瘤的诊断、治疗和预后。

如本文所用，术语“CT120蛋白”、“CT120多肽”、“肿瘤相关CT120蛋白”

白”或“肿瘤相关蛋白 CT120”可互换使用，都指具有人肿瘤相关蛋白 CT120 氨基酸序列(SEQ ID NO:2)的蛋白或多肽。该术语还包括含有或不含起始甲硫氨酸的肿瘤相关蛋白 CT120。

如本文所用，“分离的”是指物质从其原始环境中分离出来(如果是天然的物质，原始环境即是天然环境)。如活体细胞内的天然状态下的多聚核苷酸和多肽是没有分离纯化的，但同样的多聚核苷酸或多肽如从天然状态中同存在的其他物质中分开，则为分离纯化的。

如本文所用，“分离的肿瘤相关 CT120 蛋白或多肽”，“分离的 CT120 蛋白或多肽”是指肿瘤相关 CT120 蛋白多肽基本上不含天然与其相关的其它蛋白、脂类、糖类或其它物质。本领域的技术人员能用标准的蛋白质纯化技术纯化 CT120 蛋白。基本上纯的多肽在非还原聚丙烯酰胺凝胶上能产生单一的主带。CT120 蛋白多肽的纯度能用氨基酸序列分析。

本发明的多肽可以是重组多肽、天然多肽、合成多肽，优选重组多肽。本发明的多肽可以是天然纯化的产物，或是化学合成的产物，或使用重组技术从原核或真核宿主(例如，细菌、酵母、高等植物、昆虫和哺乳动物细胞)中产生。根据重组生产方案所用的宿主，本发明的多肽可以是糖基化的，或可以是非糖基化的。本发明的多肽还可包括或不包括起始的甲硫氨酸残基。

本发明还包括肿瘤相关的人 CT120 蛋白的片段、衍生物和类似物。如本文所用，术语“片段”、“衍生物”和“类似物”是指基本上保持本发明的天然肿瘤相关人 CT120 蛋白相同的生物学功能或活性的多肽。本发明的多肽片段、衍生物或类似物可以是(i)有一个或多个保守或非保守性氨基酸残基(优选保守性氨基酸残基)被取代的多肽，而这样的取代的氨基酸残基可以是也可以不是由遗传密码编码的，或(ii)在一个或多个氨基酸残基中具有取代基团的多肽，或(iii)成熟多肽与另一个化合物(比如延长多肽半衰期的化合物，例如聚乙二醇)融合所形成的多肽，或(iv)附加的氨基酸序列融合到此多肽序列而形成的多肽(如前导序列或分泌序列或用来纯化此多肽的序列或蛋白原序列)。根据本文的教导，这些片段、衍生物和类似物属于本领域熟练技术人员公知的范围。

在本发明中，术语“人肿瘤相关蛋白 CT120 多肽”或“人 NIP2 AP 蛋白多肽”可互换使用，都指具有人肿瘤相关蛋白 CT120 活性的 SEQ ID NO. 2 序列的多肽。该术语还包括具有与人肿瘤相关蛋白 CT120 相同功能的、SEQ ID NO. 2 序列的变异形式。这些变异形式包括(但并不限于)：若干个(通常为 1-50 个，较佳地 1-30 个，更佳地 1-20 个，最佳地 1-10 个)氨基酸的缺失、插入和/或取代，以及在 C 末端和/

或 N 末端添加一个或数个(通常为 20 个以内，较佳地为 10 个以内，更佳地为 5 个以内)氨基酸。例如，在本领域中，用性能相近或相似的氨基酸进行取代时，通常不会改变蛋白质的功能。又比如，在 C 末端和/或 N 末端添加一个或数个氨基酸通常也不会改变蛋白质的功能。该术语还包括人肿瘤相关蛋白 CT120 的活性片段和

5 活性衍生物。

该多肽的变异形式包括：同源序列、保守性变异数体、等位变异数体、天然突变体、诱导突变体、在高或低的严紧度条件下能与人肿瘤相关蛋白 CT120 DNA 杂交的 DNA 所编码的蛋白、以及利用抗人肿瘤相关蛋白 CT120 多肽的抗血清获得的多肽或蛋白。本发明还提供了其他多肽，如包含人肿瘤相关蛋白 CT120 多肽或其片

10 段的融合蛋白(如包含 SEQ ID NO:2 所示序列的融合蛋白)。除了几乎全长的多肽外，本发明还包括了人肿瘤相关蛋白 CT120 多肽的可溶性片段。通常，该片段具有人肿瘤相关蛋白 CT120 多肽序列的至少约 10 个连续氨基酸，通常至少约 30 个连续氨基酸，较佳地至少约 50 个连续氨基酸，更佳地至少约 80 个连续氨基酸，最佳地至少约 100 个连续氨基酸。

15 发明还提供人肿瘤相关蛋白 CT120 或多肽的类似物。这些类似物与天然人肿瘤相关蛋白 CT120 多肽的差别可以是氨基酸序列上的差异，也可以是不影响序列的修饰形式上的差异，或者兼而有之。这些多肽包括天然或诱导的遗传变异数体。诱导变异数体可以通过各种技术得到，如通过辐射或暴露于诱变剂而产生随机诱变，还可通过定点诱变法或其他已知分子生物学的技术。类似物还包括具有不同于天然

20 L-氨基酸的残基(如 D-氨基酸)的类似物，以及具有非天然存在的或合成的氨基酸(如  $\beta$ 、 $\gamma$ -氨基酸)的类似物。应理解，本发明的多肽并不限于上述例举的代表性的多肽。

25 修饰(通常不改变一级结构)形式包括：体内或体外的多肽的化学衍生形式如乙酰化或羧基化。修饰还包括糖基化，如那些在多肽的合成和加工中或进一步加工步骤中进行糖基化修饰而产生的多肽。这种修饰可以通过将多肽暴露于进行糖基化的酶(如哺乳动物的糖基化酶或去糖基化酶)而完成。修饰形式还包括具有磷酸化氨基酸残基(如磷酸酪氨酸，磷酸丝氨酸，磷酸苏氨酸)的序列。还包括被修饰从而提高了其抗蛋白水解性能或优化了溶解性能的多肽。

30 在本发明中，“人肿瘤相关蛋白 CT120 保守性变异数体”指与 SEQ ID NO: 2 的氨基酸序列相比，有至多 10 个，较佳地至多 8 个，更佳地至多 5 个，最佳地至多 3 个氨基酸被性质相似或相近的氨基酸所替换而形成多肽。这些保守性变异数体最好根据表 A 进行氨基酸替换而产生。

表 A

最初的残基	代表性的取代	优选的取代
Ala (A)	Val; Leu; Ile	Val
Arg (R)	Lys; Gln; Asn	Lys
Asn (N)	Gln; His; Lys; Arg	Gln
Asp (D)	Glu	Glu
Cys (C)	Ser	Ser
Gln (Q)	Asn	Asn
Glu (E)	Asp	Asp
Gly (G)	Pro; Ala	Ala
His (H)	Asn; Gln; Lys; Arg	Arg
Ile (I)	Leu; Val; Met; Ala; Phe	Leu
Leu (L)	Ile; Val; Met; Ala; Phe	Ile
Lys (K)	Arg; Gln; Asn	Arg
Met (M)	Leu; Phe; Ile	Leu
Phe (F)	Leu; Val; Ile; Ala; Tyr	Leu
Pro (P)	Ala	Ala
Ser (S)	Thr	Thr
Thr (T)	Ser	Ser
Trp (W)	Tyr; Phe	Tyr
Tyr (Y)	Trp; Phe; Thr; Ser	Phe
Val (V)	Ile; Leu; Met; Phe; Ala	Leu

本发明的多核苷酸可以是 DNA 形式或 RNA 形式。DNA 形式包括 cDNA、基因组 DNA 或人工合成的 DNA。DNA 可以是单链的或是双链的。DNA 可以是编码链或非编码链。编码成熟多肽的编码区序列可以与 SEQ ID NO:1 所示的编码区序列(第 91-861 位)相同或者是简并的变异体。如本文所用，“简并的变异体”在本发明中是指编码具有 SEQ ID NO:2 的蛋白质，但与 SEQ ID NO:1 所示的编码区序列有差别的核酸序列。

编码 SEQ ID NO:2 的成熟多肽的多核苷酸包括：只编码成熟多肽的编码序列；成熟多肽的编码序列和各种附加编码序列；成熟多肽的编码序列(和任选的附加编码序列)以及非编码序列。

术语“编码多肽的多核苷酸”可以是包括编码此多肽的多核苷酸，也可以是还包括附加编码和/或非编码序列的多核苷酸。

本发明还涉及上述多核苷酸的变异体，其编码与本发明有相同的氨基酸序列的多肽或多肽的片段、类似物和衍生物。此多核苷酸的变异体可以是天然发生的等

位变异体或非天然发生的变异体。这些核苷酸变异体包括取代变异体、缺失变异体和插入变异体。如本领域所知的，等位变异体是一个多核苷酸的替换形式，它可能是一个或多个核苷酸的取代、缺失或插入，但不会从实质上改变其编码的多肽的功能。

5 本发明还涉及与上述的序列杂交且两个序列之间具有至少 50%，较佳地至少 70%，更佳地至少 80%，最佳地至少 90% 相同性的多核苷酸。本发明特别涉及在严格条件下与本发明所述多核苷酸可杂交的多核苷酸。在本发明中，“严格条件”是指：(1)在较低离子强度和较高温度下的杂交和洗脱，如  $0.2 \times SSC$ , 0.1% SDS, 60°C；或(2)杂交时加有变性剂，如 50%(v/v) 甲酰胺，0.1% 小牛血清/0.1% Ficoll, 42°C 等；  
10 或(3)仅在两条序列之间的相同性至少在 95% 以上，最好是 97% 以上时才发生杂交。并且，可杂交的多核苷酸编码的多肽与 SEQ ID NO: 2 所示的成熟多肽有相同的生物学功能和活性。

本发明还涉及与上述的序列杂交的核酸片段。如本文所用，“核酸片段”的长度至少含 15 个核苷酸，较好是至少 30 个核苷酸，最好是至少 50 个核苷酸，最好是至少 100 个核苷酸以上。核酸片段可用于核酸的扩增技术(如 PCR)以确定和/或分离编码 CT120 蛋白的多聚核苷酸。  
15

本发明中的多肽和多核苷酸优选以分离的形式提供，更佳地被纯化至均质。

本发明的 DNA 序列能用几种方法获得。例如，用本领域熟知的杂交技术分离 DNA。这些技术包括但不限于：1)用探针与基因组或 cDNA 文库杂交以检出同源性核苷酸序列，和 2)表达文库的抗体筛选以检出具有共同结构特征的克隆的 DNA 片段。  
20

编码 CT120 蛋白的特异 DNA 片段序列产生也能用下列方法获得：1)从基因组 DNA 分离双链 DNA 序列；2)化学合成 DNA 序列以获得所需多肽的双链 DNA。

上述提到的方法中，分离基因组 DNA 最不常用。当需要的多肽产物的整个氨基酸序列已知时，DNA 序列的直接化学合成是经常选用的方法。如果所需的氨基酸的整个序列不清楚时，DNA 序列的直接化学合成是不可能的，选用的方法是 cDNA 序列的分离。分离感兴趣的 cDNA 的标准方法是从高表达该基因的供体细胞分离 mRNA 并进行逆转录，形成质粒或噬菌体 cDNA 文库。提取 mRNA 的方法已有多种成熟的技术，试剂盒也可从商业途径获得(Qiagene)。而构建 cDNA 文库也是通常的方法。还可得到商业供应的 cDNA 文库，如 Clontech 公司的不同 cDNA 文库。当结合使用聚合酶反应技术时，即使极少的表达产物也能克隆。  
25

可用常规方法从这些 cDNA 文库中筛选本发明的基因。这些方法包括(但不限于)：(1)DNA-DNA 或 DNA-RNA 杂交；(2)标志基因的功能出现或丧失；(3)测定 CT120 蛋白的转录本的水平；(4)通过免疫学技术或测定生物学活性，来检测基因表达的蛋白产物。上述方法可单用，也可多种方法联合应用。  
30  
35

在第(1)种方法中，杂交所用的探针是与本发明的多核苷酸的任何一部分同源，其长度至少 15 个核苷酸，较好是至少 30 个核苷酸，更好是至少 50 个核苷酸，最好是至少 100 个核苷酸。此外，探针的长度通常在 2kb 之内，较佳地为 1kb 之内。此处所用的探针通常是在本发明的基因 DNA 序列信息的基础上化学合成的 DNA 序列。本发明的基因本身或者片段当然可以用作探针。DNA 探针的标记可用放射性同位素，荧光素或酶(如碱性磷酸酶)等。

在第(4)种方法中，检测 CT120 蛋白基因表达的蛋白产物可用免疫学技术如 Western 印迹法，放射免疫沉淀法，酶联免疫吸附法(ELISA)等。

应用 PCR 技术扩增 DNA/RNA 的方法被优先用于获得本发明的基因。特别是很难从文库中得到全长的 cDNA 时，可优先使用 RACE 法(RACE-cDNA 末端快速扩增法)，用于 PCR 的引物可根据本文所公开的本发明的序列信息适当地选择，并可用常规方法合成。可用常规方法如通过凝胶电泳分离和纯化扩增的 DNA/RNA 片段。

如上所述得到的本发明的基因，或者各种 DNA 片段等的核苷酸序列的测定可用常规方法如双脱氧链终止法(Sanger et al. PNAS, 1977, 74: 5463-5467)。这类核苷酸序列测定也可用商业测序试剂盒等。为了获得全长的 cDNA 序列，测序需反复进行。有时需要测定多个克隆的 cDNA 序列，才能拼接成全长的 cDNA 序列。

本发明也涉及包含本发明的多核苷酸的载体，以及用本发明的载体或 CT120 蛋白编码序列经基因工程产生的宿主细胞，以及经重组技术产生本发明所述多肽的方法。

通过常规的重组 DNA 技术，可利用本发明的多聚核苷酸序列可用来表达或生产重组的 CT120 蛋白多肽(Science, 1984; 224: 1431)。一般来说有以下步骤：

(1).用本发明的编码肿瘤相关人 CT120 蛋白的多核苷酸(或变异体)，或用含有该多核苷酸的重组表达载体转化或转导合适的宿主细胞；

25 (2).在合适的培养基中培养的宿主细胞；

(3).从培养基或细胞中分离、纯化蛋白质。

本发明中，肿瘤相关的人 CT120 蛋白多核苷酸序列可插入到重组表达载体中。术语“重组表达载体”指本领域熟知的细菌质粒、噬菌体、酵母质粒、植物细胞病毒、哺乳动物细胞病毒如腺病毒、逆转录病毒或其他载体。在本发明中适用的载体包括但不限于：在细菌中表达的基于 T7 的表达载体(Rosenberg, et al. Gene, 1987, 56:125); 在哺乳动物细胞中表达的 pMSXND 表达载体(Lee and Nathans, J Bio Chem. 263:3521,1988)和在昆虫细胞中表达的来源于杆状病毒的载体。总之，只要能在宿主体内复制和稳定，任何质粒和载体都可以用。表达载体的一个重要特征是通常含有复制起点、启动子、标记基因和翻译控制元件。

35 本领域的技术人员熟知的方法能用于构建含 CT120 蛋白编码 DNA 序列和合适

的转录/翻译控制信号的表达载体。这些方法包括体外重组 DNA 技术、DNA 合成技术、体内重组技术等(Sambrook, et al. Molecular Cloning, a Laboratory Manual, cold Spring Harbor Laboratory. New York, 1989)。所述的 DNA 序列可有效连接到表达载体中的适当启动子上，以指导 mRNA 合成。这些启动子的代表性例子有：

5 大肠杆菌的 lac 或 trp 启动子；λ 噬菌体 PL 启动子；真核启动子包括 CMV 立即早期启动子、HSV 胸苷激酶启动子、早期和晚期 SV40 启动子、反转录病毒的 LTRs 和其他一些已知的可控制基因在原核或真核细胞或其病毒中表达的启动子。表达载体还包括翻译起始用的核糖体结合位点和转录终止子。

10 此外，表达载体优选地包含一个或多个选择性标记基因，以提供用于选择转化的宿主细胞的表型性状，如真核细胞培养用的二氢叶酸还原酶、新霉素抗性以及绿色荧光蛋白(GFP)，或用于大肠杆菌的四环素或氨苄青霉素抗性。

包含上述的适当 DNA 序列以及适当启动子或者控制序列的载体，可以用于转化适当的宿主细胞，以使其能够表达蛋白质。

15 宿主细胞可以是原核细胞，如细菌细胞；或是低等真核细胞，如酵母细胞；或是高等真核细胞，如哺乳动物细胞。代表性例子有：大肠杆菌，链霉菌属；鼠伤寒沙门氏菌的细菌细胞；真菌细胞如酵母；植物细胞；果蝇 S2 或 Sf9 的昆虫细胞；CHO、COS 或 Bowes 黑素瘤细胞的动物细胞等。

20 本发明的多核苷酸在高等真核细胞中表达时，如果在载体中插入增强子序列时将会使转录得到增强。增强子是 DNA 的顺式作用因子，通常大约有 10 到 300 个碱基对，作用于启动子以增强基因的转录。可举的例子包括在复制起始点晚期一侧的 100 到 270 个碱基对的 SV40 增强子、在复制起始点晚期一侧的多瘤增强子以及腺病毒增强子等。

本领域一般技术人员都清楚如何选择适当的载体、启动子、增强子和宿主细胞。

25 用重组 DNA 转化宿主细胞可用本领域技术人员熟知的常规技术进行。当宿主为原核生物如大肠杆菌时，能吸收 DNA 的感受态细胞可在指数生长期后收获，用  $\text{CaCl}_2$  法处理，所用的步骤在本领域众所周知。可供选择的是用  $\text{MgCl}_2$ 。如果需要，转化也可用电穿孔的方法进行。当宿主是真核生物，可选用如下的 DNA 转染方法：磷酸钙共沉淀法，常规机械方法如显微注射、电穿孔、脂质体包装等。

30 获得的转化子可以用常规方法培养，表达本发明的基因所编码的多肽。根据所用的宿主细胞，培养中所用的培养基可选自各种常规培养基。在适于宿主细胞生长的条件下进行培养。当宿主细胞生长到适当的细胞密度后，用合适的方法(如温度转换或化学诱导)诱导选择的启动子，将细胞再培养一段时间。

35 在上面的方法中的重组多肽可包被于细胞内、细胞外或在细胞膜上表达或分泌到细胞外。如果需要，可利用其物理的、化学的和其它特性通过各种分离方法分

离和纯化重组的蛋白。这些方法是本领域技术人员所熟知的。这些方法的例子包括但并不限于：常规的复性处理、用蛋白沉淀剂处理(盐析方法)、离心、渗透破菌、超处理、超离心、分子筛层析(凝胶过滤)、吸附层析、离子交换层析、高效液相层析(HPLC)和其它各种液相层析技术及这些方法的结合。

5 本发明人研究已表明，在正常肺组织中，CT120 不表达，而在发生癌变的肺细胞中，CT120 表达。因此，可通过检测 CT120 转录本或蛋白来检测肺癌。

因此，本发明重组的人肿瘤相关 CT120 蛋白或多肽有多方面的用途。这些用途包括(但不限于)：筛选促进或对抗 CT120 蛋白功能的抗体、多肽或其它配体。例如，抗体可用于抑制 CT120 蛋白的功能。用表达的重组 CT120 蛋白筛选多肽库可用于寻找有治疗价值的能抑制或刺激 CT120 蛋白功能的多肽分子。

10 本发明也提供了筛选药物以鉴定提高(激动剂)或阻遏(拮抗剂)CT120 蛋白的药剂的方法。例如，能在药物的存在下，将哺乳动物细胞或表达 CT120 蛋白的膜制剂与标记的 CT120 蛋白一起培养。然后测定药物提高或阻遏此相互作用的能力。

15 CT120 蛋白的拮抗剂包括筛选出的抗体、化合物、缺失物和类似物等。CT120 蛋白的拮抗剂可以与 CT120 蛋白结合并消除其功能，或是抑制 CT120 蛋白的产生，或是与多肽的活性位点结合使多肽不能发挥生物学功能。CT120 蛋白的拮抗剂可用于治疗用途。

本发明的多肽的拮抗剂(如反义序列和抗体)可直接用于疾病治疗，例如，各种恶性肿瘤和细胞异常增殖等，尤其是用于肺癌和肝癌的治疗。

20 本发明的多肽，及其片段、衍生物、类似物或它们的细胞可以用来作为抗原以生产抗体。这些抗体可以是多克隆或单克隆抗体。多克隆抗体可以通过将此多肽直接注射动物的方法得到。制备单克隆抗体的技术包括杂交瘤技术，三瘤技术，人 B-细胞杂交瘤技术，EBV-杂交瘤技术等。

25 可以将本发明的多肽或拮抗剂与合适的药物载体组合后使用。这些载体可以是水、葡萄糖、乙醇、盐类、缓冲液、甘油以及它们的组合。组合物包含安全有效量的多肽或拮抗剂以及不影响药物效果的载体和赋形剂。这些组合物可以作为药物用于疾病治疗。

30 本发明还提供含有一种或多种容器的药盒或试剂盒，容器中装有一种或多种本发明的药用组合物成分。与这些容器一起，可以有由制造、使用或销售药品或生物制品的政府管理机构所给出的指示性提示，该提示反映出生产、使用或销售的政府管理机构许可其在人体上施用。此外，本发明的多肽可以与其它的治疗化合物结合使用。

35 药物组合物可以以方便的方式给药，如通过局部、静脉内、腹膜内、肌内、皮下、鼻内或皮内的给药途径。CT120 蛋白以有效地治疗和/或预防具体的适应症的量来给药。施用于患者的 CT120 蛋白的量和剂量范围将取决于许多因素，如给

药方式、待治疗者的健康条件和诊断医生的判断。

CT120 蛋白的多聚核苷酸也可用于多种治疗目的。基因治疗技术可用于治疗由于 CT120 蛋白的表达异常所致的细胞增殖、发育或代谢异常。

抑制 CT120 蛋白 mRNA 的寡聚核苷酸(包括反义 RNA 和 DNA)以及核酶也在本发明的范围之内。核酶是一种能特异性分解特定 RNA 的酶样 RNA 分子，其作用机制是核酶分子与互补的靶 RNA 特异性杂交后进行核酸内切作用。反义的 RNA 和 DNA 及核酶可用已有的任何 RNA 或 DNA 合成技术获得，如固相磷酸酰胺化学合成法合成寡核苷酸的技术已广泛应用。反义 RNA 分子可通过编码该 RNA 的 DNA 序列在体外或体内转录获得。这种 DNA 序列已整合到载体的 RNA 聚合酶启动子的下游。为了增加核酸分子的稳定性，可用多种方法对其进行修饰，如增加两侧的序列长度，核糖核苷之间的连接应用磷酸硫酯键或肽键而非磷酸二酯键。

多聚核苷酸导入组织或细胞内的方法包括：将多聚核苷酸直接注入到体内组织中；或在体外通过载体(如病毒、噬菌体或质粒等)先将多聚核苷酸导入细胞中，再将细胞移植到体内等。

本发明的多肽还可用作肽谱分析，例如，多肽可用物理的、化学或酶进行特异性切割，并进行一维或二维或三维的凝胶电泳分析。

本发明还提供了针对 CT120 蛋白抗原决定簇的抗体。这些抗体包括(但不限于)：多克隆抗体、单克隆抗体、嵌合抗体、单链抗体、Fab 片段和 Fab 表达文库产生的片段。

抗 CT120 蛋白的抗体可用于免疫组织化学技术中，检测活检标本中的 CT120 蛋白。

与 CT120 蛋白结合的单克隆抗体也可用放射性同位素标记，注入体内可跟踪其位置和分布。这种放射性标记的抗体可作为一种非创伤性诊断方法用于肿瘤细胞的定位和判断是否有转移。

本发明中的抗体可用于治疗或预防与 CT120 蛋白相关的疾病。给予适当剂量的抗体可以刺激或阻断 CT120 蛋白的产生或活性。

抗体也可用于设计针对体内某一特殊部位的免疫毒素。如 CT120 蛋白高亲和性的单克隆抗体可与细菌或植物毒素(如白喉毒素，蓖麻蛋白，红豆碱等)共价结合。一种通常的方法是用巯基交联剂如 SPDP，攻击抗体的氨基，通过二硫键的交换，将毒素结合于抗体上，这种杂交抗体可用于杀灭 CT120 蛋白阳性的细胞(如表达 CT120 的肺癌细胞)。

多克隆抗体的生产可用 CT120 蛋白或多肽免疫动物，如家兔，小鼠，大鼠等。多种佐剂可用于增强免疫反应，包括但不限于弗氏佐剂等。

CT120 蛋白单克隆抗体可用杂交瘤技术生产(Kohler and Milstein. Nature, 1975, 256:495-497)。将人恒定区和非人源的可变区结合的嵌合抗体可用已有的技术生产

(Morrison et al ,PNAS,1985,81:6851)。而已有的生产单链抗体的技术(U.S. Pat No.4946778)也可用于生产抗 CT120 蛋白的单链抗体。

能与 CT120 蛋白结合的多肽分子可通过筛选由各种可能组合的氨基酸结合于固相物组成的随机多肽库而获得。筛选时，必须对 CT120 蛋白分子进行标记。

5 本发明还涉及定量和定位检测 CT120 蛋白水平的诊断试验方法。这些试验是本领域所熟知的，且包括 FISH 测定和放射免疫测定。

CT120 蛋白的多聚核苷酸可用于 CT120 蛋白相关疾病(尤其肺癌)的诊断和治疗。在诊断方面，CT120 蛋白的多聚核苷酸可用于检测 CT120 蛋白的表达与否，或检测在疾病状态下 CT120 蛋白的异常表达。而 CT120 蛋白 DNA 序列可用于对 10 活检标本的杂交以判断 CT120 蛋白的表达异常。杂交技术包括 Southern 印迹法，Northern 印迹法、原位杂交等。这些技术方法都是公开的成熟技术，相关的试剂盒都可从商业途径得到。本发明的多核苷酸的一部分或全部可作为探针固定在微阵列(Microarray)或 DNA 芯片(又称为“基因芯片”)上，用于分析组织中基因的差异表达分析和基因诊断。用 CT120 蛋白特异的引物进行 RNA-聚合酶链反应(RT-PCR) 15 体外扩增也可检测 CT120 蛋白的转录产物。

检测 CT120 蛋白基因的突变也可用于诊断 CT120 蛋白相关的疾病(尤其是肺癌)。CT120 蛋白突变的形式包括与正常野生型 CT120 蛋白 DNA 序列(如 SEQ ID NO: 1 所示的正常序列)相比的点突变、易位、缺失、重组和其它任何异常等。可用已有的技术如 Southern 印迹法、DNA 序列分析、PCR 和原位杂交检测突变。另外，突变有可能影响蛋白的表达，因此用 Northern 印迹法、Western 印迹法可间接 20 判断基因有无突变。

本发明的CT120蛋白核苷酸全长序列或其片段通常可以用PCR扩增法、重组法或人工合成的方法获得。对于PCR扩增法，可根据本发明所公开的有关核苷酸序列，尤其是开放阅读框序列来设计引物，并用市售的cDNA库或按本领域技术人员已知的常规方法所制备的cDNA库作为模板，扩增而得有关序列。当序列较长时，常常需要进行两次或多次PCR扩增，然后再将各次扩增出的片段按正确次序拼接在一起。

一旦获得了有关的序列，就可以用重组法来大批量地获得有关序列。这通常是将其克隆入载体，再转入细胞，然后通过常规方法从增殖后的宿主细胞中分离得到有关序列。

此外，还可用人工合成的方法来合成有关序列，尤其是片段长度较短时。通常，通过先合成多个小片段，然后再进行连接可获得序列很长的片段。

目前，已经可以完全通过化学合成为编码本发明蛋白(或其片段，或其衍生物)的 DNA 序列。然后可将该 DNA 序列引入本领域中的各种 DNA 分子(如载体)和细胞中。此外，还可通过化学合成将突变引入本发明蛋白序列中。

本发明首次证实了 CT120 在不同肿瘤组织有不同程度的表达，尤其在肺癌中诱导并高表达；体外 DNA 转染实验更证明 CT120 克隆对 NIH/3T3 细胞生长具有明显的促进作用。因此，CT120 是一个新的肿瘤相关基因，肿瘤的诊断、治疗及预后上具有潜在的应用价值。

下面结合具体实施例，进一步阐述本发明。应理解，这些实施例仅用于说明本发明而不用于限制本发明的范围。下列实施例中未注明具体条件的实验方法，通常按照常规条件如Sambrook等人，分子克隆：实验室手册(New York: Cold Spring Harbor Laboratory Press, 1989)中所述的条件，或按照制造厂商所建议的条件。

#### 实施例 1：PAC579 克隆中新基因的计算预测：

D17S926 位点所在的 PAC579(P579) 克隆(Genome System 公司提供)，经鸟枪法测序得到的 DNA 序列(在基康公司完成)，用 Celera 公司生物信息学系统及“Undergo”软件(Axys Pharmaceuticals)对 PAC579 基因组序列进行新基因的计算识别与预测，结果显示有一个新基因，其在 PAC579 上的位置及预测的外显子见下表：

符号	外显子编号(bp)	在 PAC579 中的位置	链
	外显子 1(122)	50808-50687	
CT120	外显子 2(169)	45607-45406	-
	外显子 3(123)	42939-42817	
	外显子 4(599)	42143-41545	

#### 实施例 2：新基因 CT120 全长 cDNA 的克隆：

用预测的外显子序列查询人 EST 数据库，根据返回的 EST 序列，可以对其进行拼接，获得一 cDNA 序列 FLJ22282(GenBank No. AK025935)。根据此序列设计引物进行 RACE 反应。

2.1 所用主要试剂：cDNA 池(Human kidney Marathon-Ready cDNAs, Clontech)，聚合酶系统(Advantage cDNA polymerase Mix, Clontech)TA 克隆系统(TOPO TA cloning)。

2.2 引物设计：用于 RACE(Rapid amplification of cDNA ends) 反应的基因特异引物应符合下列条件：(a) 长度 23-28nt；(b) GC 含量 50-70%；(c) Tm 值大

于 65°C。设计并合成以下基因特异引物引物：

120G	R	5'GTGCGACTGGCACAAAGGACAAAGAG3' (SEQ ID NO: 3)	5' RACE
120QNG	R	5'CGAATGATGACGATCCCCGAGCC3' (SEQ ID NO: 4)	5' RACE

2.3 RACE 反应：PCR 扩增反应可在 12.5μl 或 25μl 的反应体积中进行，按下列条件设置 RACE 反应：

总体积 12.5μl	
Marathon-Ready cDNAs	1.25μl
衔接引物	0.25μl
10mM dNTP	0.25μl
10×PCR 反应缓冲液	1.25μl
50×聚合酶混合物	0.25μl
H <sub>2</sub> O	9.0μl
基因特异性引物 (10pmol/μl)	0.25μl

#### PCR 反应条件：

94°C	1 分钟	1 循环
94°C	30 秒	5 循环
72°C	4 分钟	
94°C	30 秒	5 循环
70°C	4 分钟	
94°C	20 秒	25 循环
68°C	4 分钟	

5 RACE 产物的亚克隆：取回收的 PCR 产物 0.5-2.5μl，加 PCR-TOPO 载体 (Clontech 公司) 0.5μl 混匀室温放置 5 分钟置冰上，再按常规方法进行细菌转化，涂板 37°C 生长 12-16 小时，兰、白斑筛选。

#### 2.4 RACE 产物的筛选鉴定：

在 96 孔板中，每孔加入含 Amp 抗性的 LB 30μl，对于每个 RACE 反应，挑取 10-20 个白斑重组子至上述 96 孔板的 LB 中，用该菌液作模板，直接进行 PCR 反应，初步筛选出候选阳性 RACE 克隆。对候选阳性克隆进行小量液扩，抽提质粒 DNA，内切酶酶切，电泳分析，筛选出大片段 RACE 克隆，再进行 PCR 鉴定。

#### 2.5 RACE 产物的测序及序列分析：

对候选大片段阳性克隆进行测序，依据 RACE 产物的长度和该基因的 mRNA 大小，确定是否已获得该基因的全序列，全序列即包括完整的阅读框架，在第一个起始编码子 ATG 前面相同阅读框架内有终止编码子。在阅读框架的 3' 端有 polyA 序列。另外也含有相应的 5' 端和 3' 端非编码区。用 RACE 方法获得 CT120 序列和相应的编码框架，结果如 SEQ ID NO: 1-2 所示。

其中，CT120 全长 cDNA 为 2145 个碱基(SEQ ID NO:1)，其 ORF 为第 91-861 位，编码全长为 257 氨基酸的蛋白质(SEQ ID NO: 2)。

## 2.6 同源比较

同源比较结果显示 CT120 的同源物存在于不同的物种之中。CT120 在人类有两个同种型(isoform)：其中一个是由本发明蛋白 CT120A，另一个是 CT120B(AAH26023)。CT120B 比 CT120A 少第四个外显子(96 个碱基，32 个氨基酸)。在人类，还存在另一个 CT120-like 基因(NP-113666.1)。鼠中存在两个同源物 XP-133706(称之为 mCT120-like 1)和 BAB23923(mCT120-like 2)。同源比较图见图 1。其中，CT120 与 CT120B 有 223/257(86%)相同性，与 CT120-like 同源有 104/210(49%)相同性，与 mCT120-like 1 有 126/260(48%)相同性，与 mCT120-like 2 有 98/228(42%)相同性。

## 2.7 CT120 的结构分析

对 CT120 的核苷酸序列和氨基酸序列进行结构分析，发现 CT120 多肽含有以下潜在的功能域，并且具有 7 个跨膜区：

名称	SEQ ID NO:2 中的位置
PKC 磷酸化位点	39, 67, 109, 190
Casein 激酶 II 磷酸化位点	31, 61
N-肉豆蔻酰位点	148
细胞黏附序列	139-141
信号肽	1-18
跨膜区 I	4-23
跨膜区 2	42-61
跨膜区 3	76-93
跨膜区 4	113-135
跨膜区 5	145-167
跨膜区 6	179-201
跨膜区 7	216-238

## 2.6 CT120 的全长克隆:

根据 RACE 反应后所拼得的全长序列设计引物进行全长克隆，所用引物见下表。

5 120F1F: 5'CCGATGCTGCTGACGCTGGCCG3' (SEQ ID NO: 5)

120ER: 5'TGTTGGCACCAAGAAAATCCTGCTTG3' (SEQ ID NO: 6)

扩增条件用 RACE 25 $\mu$ l 反应体系及 PCR 反应条件。PCR 扩增后获得 CT120 的全长序列 1907 bp，然后装入 T-A 载体(Clontech 公司)，得到载体 CT120-T-A。

## 10 实施例 3: CT120 的多组织膜 Northern 杂交

人多组织Northern杂交膜片(MTN)购自Clontech公司，在42℃预杂交3-4小时。CT120-T-A克隆经EcoRI酶切，回收插入片段，电泳定量。取25ng DNA，加入2.5 $\mu$ l随机引物与适量水，使总体积达到13.5  $\mu$ l。煮沸5分钟，离心将液体甩至管底，加入2.5  $\mu$ l反应缓冲液，dATP、dTTP、dGTP各1  $\mu$ l，1  $\mu$ l Klenow酶，5  $\mu$ l  $^{32}$ P- $\alpha$ -dCTP。轻弹混匀，稍加离心。37℃温育20分钟，加入2  $\mu$ l 0.5M EDTA终止反应。1 ml注射器中塞入玻璃棉，加入TE饱和的Sephadex G-50。2000 rpm 5分钟。重复一次，加G-50至刻度1 ml附近。用100  $\mu$ l TE平衡三次。标记反应加75  $\mu$ l TE，上柱，离心回收。探针100℃ 5分钟变性，放至冰上。然后加入预杂交液中42℃杂交12-16小时。取膜片用1×SSC-0.05% SDS溶液42℃洗2次，每次30分钟，再用0.1×SSC-0.1% SDS 42℃洗2次，每次30分钟，最后X光片自显影。

Northern杂交结果如图2所示。CT120基因全长约为2.3kb，在心、脑、胎盘、肝、肾、胰脏、骨骼肌皆有表达，但肺中不表达。

## 25 实施例 4: 半定量反转录 PCR (RT-PCR):

本实施例通过反转录 PCR 检测 CT120 在不同的肿瘤细胞系中的表达。所用肿瘤细胞系为肺癌 SPC-A-1, 宫颈癌 C-33A, 肝癌 SMMC-7721, BEL-7402, 卵巢癌 SK-OV-3, 膀胱癌 5637, 表皮癌 A431, 乳腺癌 MCF-7。

4.1 反转录：取组织总 RNA 1  $\mu$ l 总反应体积 20 按 Superscript II RT kit (GIBCO, BRL) 操作程序合成第一链 cDNA。合成后反应体积稀释至 120, 1 $\mu$ l 约含 8ng 总 RNA, 反转录后得到的第一链 cDNA.

### 4.2 PCR 反应体系:

依次加入下列试剂

反转录第一链 cDNA 1  $\mu$ l

PCR 反应程序

	10×PCR 缓冲液	1.5 ul
	2mM dNTP	1.5 ul
	BA1 引物(上游)	1.5 ul
	BA2 引物(下游)	1.5 ul
5	CT120 F(上游)	1.5 ul
	CT120 G(下游)	1.5 ul
	Taq 酶(promega 0.5u/uL)	1 uL
	H2O	X uL
	总体积 25 uL	

10 4.3 PCR 反应程序: 94°C, 3 min; 94°C 30sec, 60°C 30 sec, 72°C 30 sec,  
26-28 循环; 72°C 5 min. PCR 反应结束后, 取 5 uL PCR 产物进行 2% 琼脂糖凝胶电泳分析.

---

$\beta$ -肌动蛋白	BA1	F	5'AAGTACTCCGTGTGGATCGG3'	SEQ ID NO: 7
	BA2	R	5'TCAAGTTGGGGACAAAAAG3'	SEQ ID NO: 8
CT120	120G	R	5'GTGCGACTGGCACAAAGGACAAAGAG3'	SEQ ID NO: 9
	120F	F	5'GGGGATCGTCATCATTGCTCCT3'	SEQ ID NO: 10

---

#### 4.3 结果

15 如图 3 所示。CT120 在肺腺癌细胞系 SPC-A-1 中表达最高; BEL-7402 和 A431 中等程度表达; C-33A, SMMC-7721, 5637, MCF-7 次之; SK-OV-3 表达较低。

鉴于 CT120 在正常肺中不表达, 在肺癌细胞中表达, 因此可通过检测 CT120 来诊断肺癌。

#### 20 实施例 5: CT120 装入真核表达载体:

选择 pcDNA4/HisMax(Invitrogen 公司)为真核表达载体, 以 cDNA 池(Clontech 公司)为模板, 用引物 120HM-F:5'ATGCTGCTGACGCTGCCGG3'(SEQ ID NO: 12); 120HM-R:5'TTAGCCATCCTTTGGCTT3'(SEQ ID NO: 13)进行扩增, 获得 CT120 的 ORF, T-A 克隆(Clontech 公司)进 pcDNA4/HisMax 真核表达载体, 获得质粒 25 pcDNA4/HisMax-CT120, 并经测序验证。挑取克隆扩增、抽质粒、酶切鉴定, 用于转化细胞。

## 实施例 6：用脂质体试剂盒转染细胞的体外实验

- 6.1 细胞株： NIH/3T3 细胞。
- 6.2 DNA： 来源于 pcDNA4/HisMax-CT120 表达质粒的 DNA。
- 6.3 脂质体： LIPOFECT AMINE™ Reagent Kit (BRL 公司)
- 5 6.4 培液： 无血清培液简称 SF-DMEM

全培液(10%小牛血清)

含 Zeocin (Invitrogen 公司) 的全培液

6 孔板(Corning 产品)。

### 6.5 DNA-脂质体复合物(DNA-liposome complex)的制备：

10 lipofectin 10 $\mu$ l 加 90 $\mu$ l SF-DMEM 混匀。DNA 1 $\mu$ g 加 100 $\mu$ l SF-DMEM 混匀。  
将稀释的 DNA 加入稀释的 lipofectin 溶液中，混匀置室温 30-45 分钟。加 0.8ml SF-DMEM 进入 DNA-lipofectin complex 中，终体积为 1.0ml。

15 6.6 转染细胞：细胞长到 50-60% 满度为好，实验前换培液一次。加 1.0ml lipofectin Reagent-DNA complex 入细胞表面，轻轻摇动，铺均匀，37℃温育 5 小时。加入 1.ml 含 20% 小牛血清 DMEM，混匀，37℃生长过夜。换培液过夜，第二天换含含 Zeocin 的全培液，常规换液筛选至克隆出现，记克隆数。

结果如图 4 所示。CT120 对 NIH/3T3 细胞生长有明显的促进作用。

20 6.7 CT120 稳定转染 NIH/3T3 细胞系的建立：挑取 CT120 稳定转染 NIH/3T3 细胞单克隆，扩大培养。单克隆细胞裂解液，12%SDS-PAGE 电泳，转膜，用抗 HisG (Invitrogen) 标签单克隆抗体检测稳定转染 NIH/3T3 细胞系中 CT120 融合蛋白的表达。

结果如图 5 所示。在所测试的 6 个克隆中 5 个克隆有 CT120 的表达，分子量约为 34KDa。

## 25 实施例 7：免疫组织化学检测 CT120 在肺癌及癌旁组织中的表达

7.1 兔抗 CT120 蛋白多克隆抗体的制备：用肽合成仪 (Applied Biosystem 公司) 合成 CT120 蛋白的 C-端 15 肽氨基酸序列 CRKAVRLFDTPQAKK (SEQ ID NO: 11) 的寡肽，用 Maleimide Activated BSA, KLH 偶联试剂盒 (Sigma) 把合成的多肽偶联到 KLH 上，然后免疫新西兰大白兔，制备兔抗 CT120 多克隆抗体。

30 7.2 免疫组织化学检测 CT120 在肺癌及癌旁组织中的表达：肺癌及癌旁组织取自肺癌患者临床手术切除组织。用于免疫组织化学检测的肺癌和癌旁肺组织

标本用 10% 中性缓冲福尔马林固定，石蜡包埋， $5 \mu\text{m}$  厚切片，用兔抗 CT120 多克隆抗体 (1: 150 稀释) 作为第一抗体，应用 Envision System 两步法检测 Kit (mouse)，DAB 显色，Mayer 氏苏木素复染核。

结果如图 6 所示。CT120 基因在肺癌组织的癌细胞中高表达 (++)，而在癌旁肺组织中几乎不表达 (--)。

在本发明提及的所有文献都在本申请中引用作为参考，就如同每一篇文献被单独引用作为参考那样。此外应理解，在阅读了本发明的上述讲授内容之后，本领域技术人员可以对本发明作各种改动或修改，这些等价形式同样落于本申请所附权利要求书所限定的范围。

### 序列表

<110> 上海市肿瘤研究所

<120> 人类 17p13.3 区域内人肿瘤相关基因 CT120 及其编码蛋白

<130> 024832

<160> 13

<170> PatentIn version 3.1

<210> 1

<211> 2145

<212> DNA

<213> 智人(Homo sapiens)

<220>

<221> CDS

<222> (91)..(861)

<223>

<400> 1

cggagggttg aaatcgccgcg gccgggcccgg ggccgcgccga gcccacccca gcccacgcggc 60

gccagcgagg cggccggacc cgcaaaaaatgg atg ctg ctg acg ctg gcc ggg ggc 114  
Met Leu Leu Thr Leu Ala Gly Gly  
1 5

gcg ctc ttc ttc ccg ggg ctc ttc gcg ctc tgc acc tgg gcg ctg cgc 162  
Ala Leu Phe Phe Pro Gly Leu Phe Ala Leu Cys Thr Trp Ala Leu Arg  
10 15 20

cac tcc cag ccc gga tgg agc cgc acc gac tgc gtg atg atc agc acc 210  
His Ser Gln Pro Gly Trp Ser Arg Thr Asp Cys Val Met Ile Ser Thr  
25 30 35 40

agg ctg gtt tcc tcg gtg cac gcc gtg ctg gcc acc ggc tcg ggg atc 258  
Arg Leu Val Ser Ser Val His Ala Val Leu Ala Thr Gly Ser Gly Ile  
45 50 55

gtc atc att cgc tcc tgc gac gac gtg atc acc ggc agg cac tgg ctt 306  
Val Ile Ile Arg Ser Cys Asp Asp Val Ile Thr Gly Arg His Trp Leu  
60 65 70

gcc cgg gaa tat gtg tgg ttt ctg att cca tac atg atc tat gac tcg 354  
Ala Arg Glu Tyr Val Trp Phe Leu Ile Pro Tyr Met Ile Tyr Asp Ser  
75 80 85

tac gcc atg tac ctc tgt gaa tgg tgc cga acc aga gac cag aac cgt      402  
 Tyr Ala Met Tyr Leu Cys Glu Trp Cys Arg Thr Arg Asp Gln Asn Arg  
 90                        95                        100

gcg ccc tcc ctc act ctt cga aac ttc cta agt cga aac cgc ctc atg      450  
 Ala Pro Ser Leu Thr Leu Arg Asn Phe Leu Ser Arg Asn Arg Leu Met  
 105                        110                        115                        120

atc aca cat cat gcg gtc att ctc ctt gtc ctt gtg cca gtc gca cag      498  
 Ile Thr His His Ala Val Ile Leu Leu Val Leu Val Pro Val Ala Gln  
 125                        130                        135

agg ctc cgg gga gac ctt ggg gac ttc ttt gtc ggc tgc atc ttc acg      546  
 Arg Leu Arg Gly Asp Leu Gly Asp Phe Phe Val Gly Cys Ile Phe Thr  
 140                        145                        150

gca gaa ctg agc act ccg ttt gtg tcg ctg ggc agg gtt ctg att cag      594  
 Ala Glu Leu Ser Thr Pro Phe Val Ser Leu Gly Arg Val Leu Ile Gln  
 155                        160                        165

cta aag cag cag cac acc ctt ctg tac aag gtg aat gga atc ctc acg      642  
 Leu Lys Gln Gln His Thr Leu Leu Tyr Lys Val Asn Gly Ile Leu Thr  
 170                        175                        180

ctg gcc acc ttc ctt tcc tgc cgg atc ctt ctc ttc ccc ttc atg tac      690  
 Leu Ala Thr Phe Leu Ser Cys Arg Ile Leu Leu Phe Pro Phe Met Tyr  
 185                        190                        195                        200

tgg tcc tat ggc cgc cag cag gga cta agc ctg ctc caa gta ccc ttc      738  
 Trp Ser Tyr Gly Arg Gln Gln Gly Leu Ser Leu Leu Gln Val Pro Phe  
 205                        210                        215

agc atc cca ttc tac tgc aac gtg gcc aat gcc ttc ctc gta gct cct      786  
 Ser Ile Pro Phe Tyr Cys Asn Val Ala Asn Ala Phe Leu Val Ala Pro  
 220                        225                        230

cag atc tac tgg ttc tgt ctg ctg tgc agg aag gca gtc cgg ctc ttt      834  
 Gln Ile Tyr Trp Phe Cys Leu Leu Cys Arg Lys Ala Val Arg Leu Phe  
 235                        240                        245

gac act ccc caa gcc aaa aag gat ggc taaaatgctcc tgggagtcag      881  
 Asp Thr Pro Gln Ala Lys Lys Asp Gly  
 250                        255

gcgcagcctc acaccagctg cctccctccac tcagcattcc atggaccaaa ttgtgccctg      941

ggtagcctca gactttgggt attgataagc cgatggattt gagttttct aaagaatatt      1001

catattacct ccttttcta acttgcccta tttgcaaacg cactttgt aataacaacta 1061  
 ttgggtcctg tcagacctcc acggacagca aagtggttt aatgcaagcc caaggatcct 1121  
 tcttaaggic ttatctcaag agctctggga ggtggaagca tggggggga tcggtgacc 1181  
 agggtggtaa gtgtctgcac atctgcctgt ccctgtatca gcggctaccc accttccaaa 1241  
 ccactcagga cagtaccgt ggcactggc ccgcagaagc aaggatgac ttggttctg 1301  
 gaagtaatgt cgtcttgta cattggcctg ggacaatcat tgtggtagg tagttattga 1361  
 tcgtttacta gataacccat tggttcttg cctcatcctc tcatccatgg gtcagagttg 1421  
 aattcttatg tctatagact tccaaatcaga agtctcactg gtggggctgg ggggggggc 1481  
 aggcaggagg catggatggg aacctgagta ggtagtgtgg ccaagagatc agcacaacct 1541  
 ttgcaggctg acttgctaag tctgacagtg acaaacttgt gagcttactg cagtcagtca 1601  
 cagaggctgt tcttttac acaccccttc atgcccggct ttccccatat ccacatgcag 1661  
 agggcgagct cataaaacta cagggaaagcg tgaaatgatg gctttggtag ctgtttactg 1721  
 ggtaacccca ctgtgacact gtcctttca tgtgatgtgg aaacctactt ctgtcctcca 1781  
 aaccatgaaa tgtgtcatct agactgcaga gtactcgagt gctttgcctc ccgatatgcc 1841  
 agagcttgtg gtccaaagcc cattcctgtg tgtccgtcct gccathtagc cacagaaggc 1901  
 tgcggagtga ggcggcagct agcctggcca gtggctgtcc cgtggaccga cacctgcgcc 1961  
 cccttctgca agcaggattt tctggtgcca acactcattc atcattcccg atcaactagg 2021  
 atgaatttaa gactgtgcta ccatgtgtc tcaagtggta gttaaaaag tggattttta 2081  
 aagtgcctt caattgtctg tgaacgtcta aaggactgat ttgtctaaaaaaa 2141  
 aaaa 2145

<210> 2  
 <211> 257  
 <212> PRT  
 <213> 智人(Homo sapiens)

<400> 2

Met Leu Leu Thr Leu Ala Gly Gly Ala Leu Phe Phe Pro Gly Leu Phe

1

5

10

15

Ala Leu Cys Thr Trp Ala Leu Arg His Ser Gln Pro Gly Trp Ser Arg  
20                           25                           30

Thr Asp Cys Val Met Ile Ser Thr Arg Leu Val Ser Ser Val His Ala  
35                           40                           45

Val Leu Ala Thr Gly Ser Gly Ile Val Ile Ile Arg Ser Cys Asp Asp  
50                           55                           60

Val Ile Thr Gly Arg His Trp Leu Ala Arg Glu Tyr Val Trp Phe Leu  
65                           70                           75                           80

Ile Pro Tyr Met Ile Tyr Asp Ser Tyr Ala Met Tyr Leu Cys Glu Trp  
85                           90                           95

Cys Arg Thr Arg Asp Gln Asn Arg Ala Pro Ser Leu Thr Leu Arg Asn  
100                           105                           110

Phe Leu Ser Arg Asn Arg Leu Met Ile Thr His His Ala Val Ile Leu  
115                           120                           125

Leu Val Leu Val Pro Val Ala Gln Arg Leu Arg Gly Asp Leu Gly Asp  
130                           135                           140

Phe Phe Val Gly Cys Ile Phe Thr Ala Glu Leu Ser Thr Pro Phe Val  
145                           150                           155                           160

Ser Leu Gly Arg Val Leu Ile Gln Leu Lys Gln Gln His Thr Leu Leu  
165                           170                           175

Tyr Lys Val Asn Gly Ile Leu Thr Leu Ala Thr Phe Leu Ser Cys Arg  
180                           185                           190

Ile Leu Leu Phe Pro Phe Met Tyr Trp Ser Tyr Gly Arg Gln Gln Gly  
195                           200                           205

Leu Ser Leu Leu Gln Val Pro Phe Ser Ile Pro Phe Tyr Cys Asn Val  
 210 215 220

Ala Asn Ala Phe Leu Val Ala Pro Gln Ile Tyr Trp Phe Cys Leu Leu  
 225 230 235 240

Cys Arg Lys Ala Val Arg Leu Phe Asp Thr Pro Gln Ala Lys Lys Asp  
 245 250 255

Gly

<210> 3  
 <211> 25  
 <212> DNA  
 <213> 人工序列

<220>  
 <221> misc\_feature  
 <223> 引物

<400> 3  
 gtgcgactgg cacaaggaca aagag

25

<210> 4  
 <211> 23  
 <212> DNA  
 <213> 人工序列

<220>  
 <221> misc\_feature  
 <223> 引物

<400> 4  
 cgaatgatga cgatccccga gcc

23

<210> 5  
 <211> 22  
 <212> DNA  
 <213> 人工序列

<220>  
<221> misc\_feature  
<223> 引物  
  
<400> 5  
ccgatgctgc tgacgctggc cg

22

<210> 6  
<211> 25  
<212> DNA  
<213> 人工序列  
  
<220>  
<221> misc\_feature  
<223> 引物  
  
<400> 6  
tgttggcacc agaaaatcct gcttg

25

<210> 7  
<211> 20  
<212> DNA  
<213> 人工序列  
  
<220>  
<221> misc\_feature  
<223> 引物  
  
<400> 7  
aagtactccg tgtggatcgg

20

<210> 8  
<211> 20  
<212> DNA  
<213> 人工序列  
  
<220>  
<221> misc\_feature  
<223> 引物  
  
<400> 8  
tcaagttggg ggacaaaaag

20

<210> 9  
<211> 25

<212> DNA  
 <213> 人工序列

<220>  
 <221> misc\_feature  
 <223> 引物

<400> 9  
 gtgcgactgg cacaaggaca aagag

25

<210> 10  
 <211> 23  
 <212> DNA  
 <213> 人工序列

<220>  
 <221> misc\_feature  
 <223> 引物

<400> 10  
 ggggatcgtc atcattcgct cct

23

<210> 11  
 <211> 15  
 <212> PRT  
 <213> 人工序列

<220>  
 <221> MISC\_FEATURE  
 <222> (1)..(15)  
 <223> 对应于 CT120 蛋白 C-端的寡肽

<400> 11

Cys Arg Lys Ala Val Arg Leu Phe Asp Thr Pro Gln Ala Lys Lys  
 1 5 10 15

<210> 12  
 <211> 20  
 <212> DNA  
 <213> 人工序列

<220>  
 <221> misc\_feature  
 <223> 引物

<400> 12

atgctgctga cgctggccgg

20

<210> 13

<211> 20

<212> DNA

<213> 人工序列

<220>

<221> misc\_feature

<223> 引物

<400> 13

ttagccatcc ttttggctt

20

02-10-03

32

# 说 明 书 附 图

CT120		
CT120B	MLTPMVAGGVVFPGFLLSKNTLQRLPQ	28
CT120-like	LAIPSSPPPSLNLAFLSLLDPLVSLPGFKSPCLPQWWLGGWCSPDSSYPRTRSRGCPS	60
mCT120-like 1	MLLTLAGGALFFPGLFALCTWALRRS-Q	27
mCT120-like 2	MLLTLAGGALFFPGLFALCTWALRRS-Q	27
CT120	MASTAGYIVSTSCKHIIDDDQHHLSSAYTQFAVPYFIYD	38
CT120B	LRWEEADAVIVSARLVSSVQAIMASTAGYIVSTSCKHIIDDDQHHLSSAYTQFAVPYFIYD	88
CT120-like	CAGREADAVIVSARLVSSVQAIMASTAGYIVSTSCKHIIDDDQHHLSSAYTQFAVPYFIYD	120
mCT120-like 1	PGWSRTDCVMISTRLVSSVHAVLATGSGIVIIRSCDDVITGRHWLAREYVWFILIPYMIYD	87
mCT120-like 2	PGWSRTDCVMISTRLVSSVHAVLATGSGIVIIRSCDDVITGRHWLAREYVWFILIPYMIYD	87
	... :: :: : * : * : * . : *** : * . * : ** : ***	
CT120	IYAMFLCHWHKHQVKHGDDGAARAPGSTWAIARGYLKEFLMVLHHAAMVLVCFPLSV	98
CT120B	IYAMFLCHWHKHQVKHGGEDEGTPRALGSTWAVVRGYLKEFLMVLHHAAMVLVCFPLSV	148
CT120-like	IYAMFLCHWHKHQVKHGGEDEGTPRALGSTWAVVRGYLKEFLMVLHHAAMVLVCFPLSV	180
mCT120-like 1	SYAMYLCCEWCRTRDQ—NRAPS—LTLRNFLSRNRLMITHHAVILFVLVPVAQ	136
mCT120-like 2	SYAMYLCCEWCRTRDQ—NRAPS—LTLRNFLSRNRLMITHHAVILFVLVPVAQ	136
	***:***.* : : : ** . * : * : ***. : : : * . : :	
CT120	VWRQKGDFFLGCMMAEVSTPFVCLGKILIQQHTLLHKVNNGALMLLSFLCCRVLIF	158
CT120B	VWRQKGDFFLGCMMAEVSTPFVCLGKILIQQHTLLHKVNNGALMLLSFLCCRVLIF	208
CT120-like	VWRQKGDFFLGCMMAEVSTPFVCLGKILIQQHTLLHKVNNGALMLLSFLCCRVLIF	240
mCT120-like 1	RLRGDLGDFVGCIFTAELSTPFVSLGRVLQLQQHTLLYKVNGILTLATFLSCRILLF	196
mCT120-like 2	—LKQQHTLLYKVNGILTLATFLSCRILLF	164
	*****:**** * * :**,**:***	
CT120	PYLYWAYGRHAGLPLLAVPLAIPAHVNLAQQLYWFPLICRGACRLFWPRS——	214
CT120B	PYLYWAYGRHAGLPLLSPVMAIPAHVNLAQQLYWFPLICRGACRLFRPRC——	264
CT120-like	PTCTGPTGATLACPCSQCPWPS-CATSTWARTAPRTQLYWLSTMCRGDCGLFRPRAPTHP	299
mCT120-like 1	PFMYWSYGRQQGLSLLQVFSIPFYCNVANAFLVAPQIYWFCLLCRKAVRLFDTQ——	252
mCT120-like 2	PFMYWSYGRQQGLSLLQVFSIPFYCNVANAFLVAPQIYWFCLLCRKAVRLFDTQ——	220
	* . * . . * . . . *;*: *;*: ** .	
CT120	RPPP-ACQAQD	224
CT120B	SPPPSPCQTQD	275
CT120-like	LLVRPRTEARPWNPPPPPAPVETVHGNQCVSWGGGDESQKSLSLTAPRQMDLE	363
mCT120-like 1	AKKDG	257
mCT120-like 2	AKKDG	225

图 1

00-10-003

32

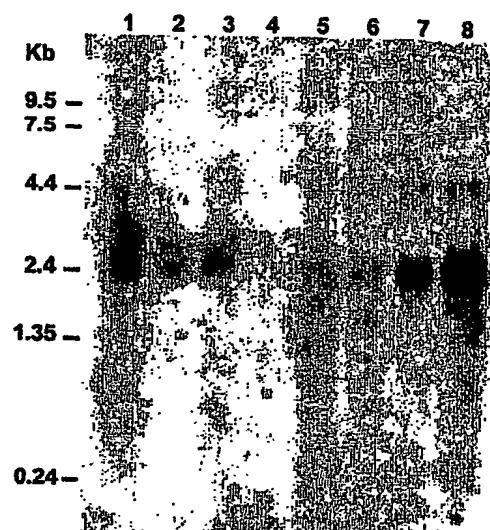


图 2

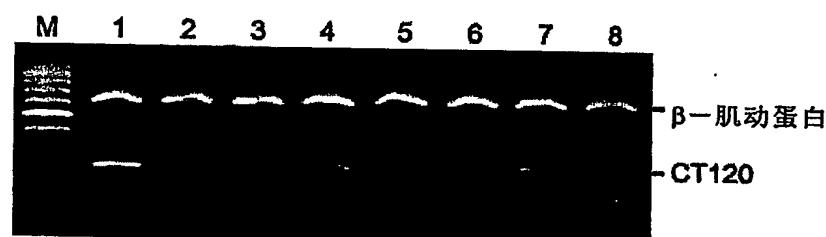


图 3

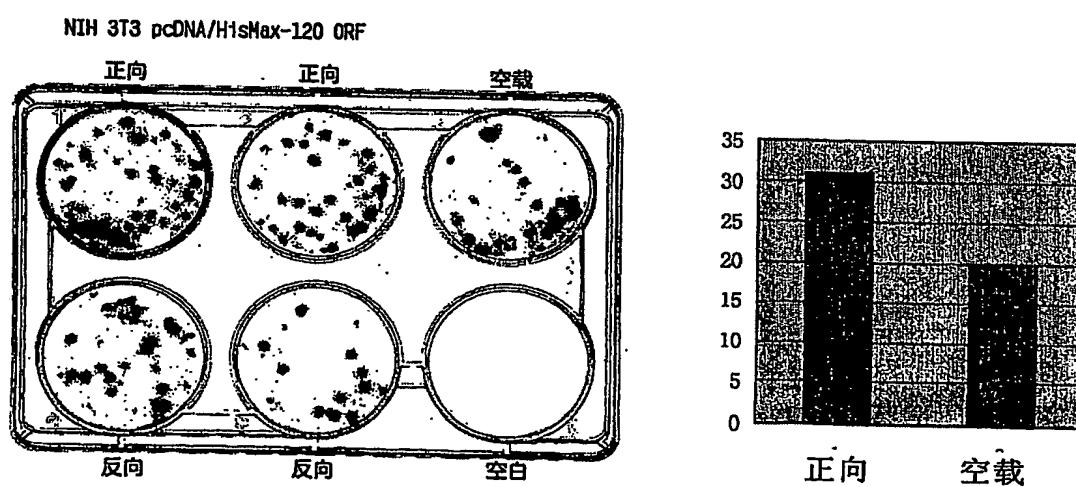


图 4

00-10-03

34

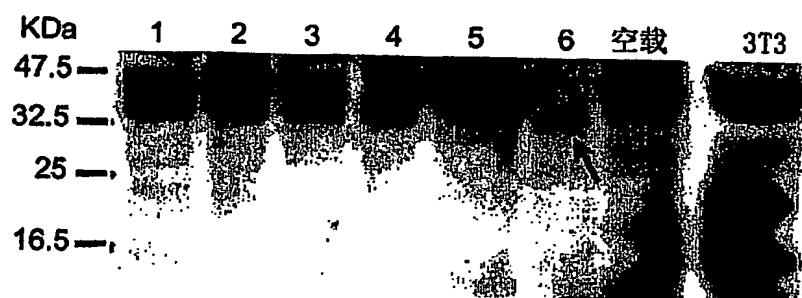
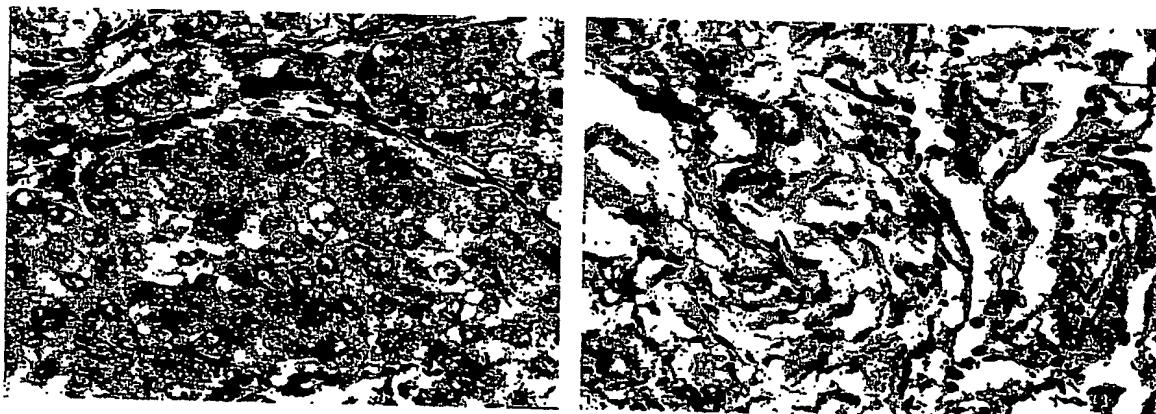


图 5



A 肺癌组织

B 肺癌癌旁组织

图 6

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**